

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
21 June 2001 (21.06.2001)

PCT

(10) International Publication Number  
**WO 01/45088 A1**

(51) International Patent Classification<sup>7</sup>: **G10L 13/08**,  
15/26, 21/06

34222 (US). YUHAS, Mark, J.; 2904 Homasassa Road,  
Sarasota, FL 34239 (US). DORETY, Michael, F.; 1963  
Rolling Green Circle, Sarasota, FL 34240 (US).

(21) International Application Number: PCT/US00/32452

(74) Agents: **BLAKE, William, A.**, et al.; Jones, Tullar &  
Cooper, P.C., P.O. Box 2266 Eads Station, Arlington, VA  
22202 (US).

(22) International Filing Date:  
12 December 2000 (12.12.2000)

(25) Filing Language: English

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,  
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,  
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,  
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,  
TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(26) Publication Language: English

(30) Priority Data:  
60/171,106 16 December 1999 (16.12.1999) US  
09/610,675 7 July 2000 (07.07.2000) US

(71) Applicant: **INTERACTIVE SOLUTIONS, INC.**  
[US/US]; 2150 Whitfield Industrial Way, Sarasota, FL  
34243-4046 (US).

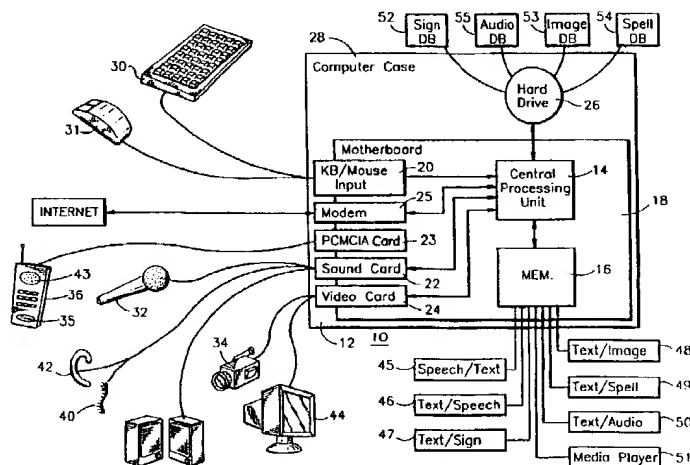
(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors: **GREENE, Morgan, Jr.**; 6628 Case Ave.,  
Bradenton, FL 34207-5620 (US). **GREENE, Virginia**;  
6628 Case Ave., Bradenton, FL 34207-5620 (US). **NEW-**  
**MAN, Harry, E.**; 351 Tahitian Drive, Ellenton, FL

Published:  
— With international search report.

[Continued on next page]

(54) Title: ELECTRONIC TRANSLATOR FOR ASSISTING COMMUNICATIONS



(57) Abstract: Speech data input to an electronic translator are simultaneously delivered to a user as multiple streams, in synchronized audible, visual and text formats. Preferably, the system converts input speech (32) to text (45), and then translates the text to any of sign language (49), animation (48), and computer-generated speech (50). The sign language and animation translations preferably use digital movies of a person signing words, phrases and finger-spelled words, and animations of the words, selectively accessed from databases (52-55) and displayed (44). Additionally, computer-generated speech is input to various hearing-enhancement devices such as cochlear implants (40) and hearing aids (42), or output to devices such as speakers. A high-speed personal computer (28) generates the text, video-signing, and audible streams simultaneously in real time. The synchronized data streams are presented concurrently, so that mental comprehension can occur. The translator can also interface with other communications devices, such as telephones (36). A hearing-impaired person is able to use a keyboard (30) or mouse (31) to converse or respond.

WO 01/45088 A1

**WO 01/45088 A1**

---



*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**ELECTRONIC TRANSLATOR FOR ASSISTING COMMUNICATIONS****BACKGROUND OF THE INVENTION****Cross Reference To Related Applications**

This application claims the benefit under 35 USC 119(e) of U.S. Provisional  
5 Application No. 60/171,106, filed December 16, 1999.

**Field of the Invention**

The present invention relates in general to an electronic translator system and methods  
that are particularly suited for facilitating two-way conversations or communications between  
speaking and/or hearing disabled individuals, and a person capable of normal  
10 communications.

**Description of the Prior Art**

In the United States, more than 28 million people, or about 10 percent of the  
population, have hearing impairments. These individuals often go through life without the  
opportunity for education, employment or the ability to lead normal lives. In the past 25  
15 years there have been very few technological advancements related to the improvement of  
education or the communication skills of the deaf, hearing-impaired, blind deaf or hearing  
mute. Although various types of hearing aids including Cochlear implants and external FM  
devices have brought some degree of relief to the more profound hearing-impaired, these  
devices suffer from a number of shortcomings that impede their ability to assist the hearing-  
20 impaired individual in improving their communication skills.

For example, cochlear implants are intended to act as artificial cochleae by simulating  
the sensory hairs that a deaf person is lacking. In response to sensed sound that may come  
from a person's voice, the cochlear implant generates electrical impulses that stimulate the  
hearing receptors in the deaf person's brain. However, the devices respond not only to a  
25 person's voice, but also to all types of ambient background noise, including other people's  
voices, for example. As a result, the cochlear implants generate signals from numerous aural  
sources that the hearing impaired individual cannot distinguish from each other. A typical  
outcome is that a hearing impaired individual with a cochlear implant will simply turn the  
device off since they cannot associate the inputs generated by the implants with words spoken  
30 by a person with whom they are trying to communicate.

Other electronic devices have been developed recently that are designed to assist  
hearing impaired individuals in learning sign language, for example. These types of devices  
typically receive as input, one or more phrase or words that are entered with a keyboard, for

example, and generate animations of signing hand movements that are associated with the entered words and phrase. These types of devices are certainly useful learning tools for the hearing impaired individual, but are of little value for assisting real-time communications. To date, known real-time communication aids for the hearing impaired have been limited  
5 either to the aforementioned cochlear implants and other electronic hearing aids, and devices that can convert spoken words to printed text. Speech-to-text translators, while useful for communicating with a hearing impaired individual, cannot be used to teach the hearing impaired individual the sound of the translated words as the text is generated. Until now, a need has therefore remained for a communications and learning aid for hearing impaired and  
10 other individuals that does not suffer from the shortcomings of known devices.

### SUMMARY OF THE INVENTION

The present invention addresses the shortcomings of known communication and learning assistance devices for hearing impaired individuals, or the like, through provision of an electronic translator system and methods that translate input speech and text into multiple  
15 streams of data that are simultaneously delivered to a user. Preferably, the data is delivered in audible, visual and, in the case of speech input, text formats. These multiple data streams are delivered to the hearing-impaired individual in a synchronized fashion, thereby creating a cognitive response. In this manner, the present invention allows a deaf or hearing-impaired person to comprehend spoken language, to achieve two-way communications without the  
20 requirement of a human translator, to learn natural speech, to learn sign language and to learn to read. The invention is also capable of interconnection with cochlear implants and hearing aids, and alleviates the data overload caused by multiple sounds with varying volume levels being introduced to the human brain simultaneously.

To achieve the foregoing functionality, the system of the present invention, in its most  
25 preferred form, is implemented using a personal computer and various conventional peripheral devices. These include, for example, a video display monitor, keyboard, mouse, one or more audio input devices, e.g., microphone, telephone transmitter, etc. and one or more audio output devices, e.g., audio speakers, telephone receiver and hearing enhancement devices used by the deaf or hearing-impaired, such as cochlear implants and hearing aids.  
30 The computer is loaded with a number of software programs or modules that perform various translations on received speech or text. In particular, speech received by the audio input device is preferably converted to text, sign language, computer-generated speech for input to the audio devices, and animation or still images that are associated with the words (e.g., an

image of an object). The sign language translation is preferably implemented by using the medium of digital movie images in which videos of a person signing words, phrase and finger spelled words are selectively accessed from a database and displayed on the video display monitor. A database of animations and images is also provided, with each of the animations or images being associated with a word or phrase in the text stream. Other databases that are also preferably provided include a phonetic spelling database to enable the phonetic spelling of words to be displayed, and an audio sound database that contains brief audio segments of environmental sounds that are associated with the words or phrase in the text stream.

The computer is preferably selected to have a fast enough processor, e.g., 500 MHz or higher, that the text, signing, audio, animation and other streams can be generated virtually simultaneously in real time, and synchronized with one another. Once synchronized, the data streams are presented to the user concurrently in a method that allows the process of mental comprehension to occur. More particularly, as the spoken words are translated into text, or the words are initially entered as text, they are simultaneously displayed on the video display monitor along with the corresponding signing videos and animation or images, and are supplied to the audio output devices. Preferably, the text words are highlighted on the screen as they are signed and spoken through the audio devices. In this manner, the user of the translator can readily associate sounds generated by a cochlear implant or a hearing aid with the signed and printed words, thus greatly improving the learning and comprehension process. In addition, the problem associated with background noise interference with the cochlear implant or hearing aid is eliminated since the generated sounds correspond only to the words of text.

Preferably, the hearing-impaired person is also able to use the system's keyboard or mouse to freely converse or respond. Words and phrases typed into the system are converted to computer-generated voice and delivered to the intended audience by the audio speakers or other means, such as the standard or cellular telephone. At this point the system has facilitated two-way conversation between a person capable of normal communications and a hearing-impaired, deaf or mute individual without the requirement of a human translator.

In addition to facilitating two-way conversation, the present invention also enables the deaf and mute to learn to speak. If an individual is physically capable of vocalizing, the invention can be used as a tool to teach linguistics. The user can utilize the computer's keyboard or mouse to generate words or phrases. The user can then practice repetition of those words or phrases while at the same time hearing a computer generated simulation of

their own voice, thereby allowing them the ability to practice and improve speech. Once the user becomes proficient at basic speech, they can use the system's microphone capability in place of the keyboard or mouse. When the system recognizes the user's speech patterns it will present them in a synchronized fashion in text, sign language and by computer-generated voice. The generation of these three data streams will support the user's comprehension of the clarity in which he is speaking thereby affording the opportunity for improvement.

The present invention can also be employed to teach the deaf, hearing-impaired, and even individuals with normal hearing and speech skills, to comprehend, communicate and translate multiple variations of sign language. For example, these include but are not limited to American Sign Language, Signed Exact English (SEE), pidgin and various forms of international sign language. As is well known, sign language is a system whereby the hands and arms are used to motion or gesture a communication. For many of the deaf and hearing-impaired, it is the first language learned and crucial to their ability to communicate and achieve an education. The present invention allows sign language to be taught in conjunction with human intervention as in the traditional teacher student scenario, or in a self teach mode whereby the subject can either speak or use the system's keyboard or mouse to achieve repetitious comprehension of motions and gestures required to create a word or sentence.

Yet another preferred feature of the invention is enabling the deaf, hearing impaired or hearing mute, to make and receive standard or cellular telephone calls. To accomplish this task, the system functions as previously described, but allows the option of dialing telephone numbers from the keyboard or with use of the mouse. Once the computer has dialed the number, a digital signal is transmitted through the system that announces the fact that the call is being generated by a hearing or speech impaired individual. The called party will respond either by voice or telephone keypad to acknowledge receipt of the call. The system then converts the voice or digital response to computer generated voice, text and sign language. In this way the deaf, hearing impaired or hearing mute individual is able to freely converse by use of the system and its keyboard or mouse. This feature is critical in that it allows the impaired individual to communicate with emergency 911 services. The optional use of a standard telephone coupling device is also provided.

#### BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other features and advantages of the present invention will become apparent from the following detailed consideration of a preferred embodiment thereof, taken in conjunction with the accompanying drawings, in which:

FIG. 1 is a schematic block diagram of an electronic translator that is constructed in accordance with a preferred embodiment of the present invention, and show the translator's hardware elements;

FIG. 2 is an illustration of a sample display screen showing the various windows that can be simultaneously displayed on a video display device in the preferred embodiment;

FIG. 3A is an illustration of a signing window that forms part of the display screen;

FIG. 3B is an illustration of a drop down menu that is employed for selecting signing display options;

FIG. 4 is an illustration of a control window that forms part of the display screen;

FIG. 5 is an illustration of a the various windows of the display screen showing how the windows can be selectively docked to one another or unlocked from each other; and

FIGs. 6A-6H are flowcharts illustrating the software flow of the preferred embodiment of the present invention.

#### DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

FIG. 1 illustrates the hardware of an electronic translator 10 that is constructed in accordance with a preferred embodiment of the present invention. The hardware includes a personal computer 12 comprised of a central processing unit (CPU) 14 and one or more memory chips 16 mounted on a motherboard 18. Preferably, the CPU has a clock speed of 500 MHz or greater to insure that multiple translations of speech or text data can be made simultaneously in real time, although a processor having a somewhat slower clock speed could also be used. Also mounted on the motherboard 18 are a keyboard/mouse input circuit 20, a sound card 22, a PCMCIA card 23, a video card 24 and a modem 25 or other communications interface for connecting the translator 10 to the Internet, for example. The CPU 14 is also interfaced to a hard drive 26. As is also conventional, all of the foregoing elements are preferably housed in a housing 28, such as a conventional PC desktop or tower housing, or a laptop housing.

The PC 12 receives multiple inputs from a number of peripheral devices, including a keyboard 30 and a mouse 31 via the keyboard/mouse input circuit 20, a microphone 32 via the sound card 22, a video camera 34 via the video card 24 and a transmitter 35 in a standard or cellular telephone 36 via the PCMCIA card 23, although the sound card 22 could also be used for this purpose. The sound card 22 also supplies output signals to a number of audio devices, including one or more speakers 38, one or more assisted learning devices including a cochlear implant 40 and a hearing aid 42, while the PCMCIA card 23 supplies audio signals

to a receiver 43 in the telephone 36 (again, the sound card 22 could also be used for this purpose). A left channel in the sound card 22 is employed for sending output signals to the telephone 36, cochlear implant 40 and hearing aid 42, while a right channel, that can be selectively turned on or off, is employed for sending output signals to the speakers 38. The video card 24 supplies output video signals to a display device, such as a video display monitor 44.

As will be discussed in greater detail in conjunction with the flowcharts of FIGs. 6A-6H, the electronic translator 10 receives speech inputs from an individual through the microphone 32, and converts these speech inputs into multiple forms. These include text, signing and other images that are displayed on the monitor 44, and audio signals that are supplied to one of the audio output devices, including the speakers 38, cochlear implant 40, hearing aid 42 and/or telephone 36. In addition, an individual using the translator 10 can input text via the keyboard 30, or access text from the Internet via the modem 25, for display on the video display 44, or translation into synthesized speech that is played on the speakers 38 or transmitted through the telephone 36.

To provide the various translation functions, the PC 12 is preferably programmed with a number of software programs or modules, including a speech-to-text translator 45, a text-to-speech translator 46, a text-to-sign language translator 47, a text-to-animation or image translator 48, a text-to-phonetic spelling translator 49, a text-to-audio sound translator 50 and a media player 51 for playing signing videos and audio sounds. Preferably, a number of these programs are commercially available off-the-shelf products. For example, the speech-to-text translator 45 can be Dragon System's NATURALLY SPEAKING or IBM's VIA VOICE, while the text-to-speech translator 46 can be MICROSOFT's text-to-speech engine, and the media player 51 is preferably WINDOWS MEDIA PLAYER.

A number of databases are stored in the hard drive 26 that are employed by the various translators. The text-to-sign language translator 47 works in a unique manner by accessing video movies of signed words and letters that are stored in a signing database 52. The database 52 contains thousands of short (e.g., 2 seconds) video clips, each of which shows a person signing a particular word. The clips are specially produced so that many of them can be displayed seamlessly in sequence as translated sentences are signed. This is accomplished by insuring that the person signing the words begins and ends each video clip with their hands in the same folded position as illustrated in FIG. 3A, for example. The text-to-sign language translator 47 functions by comparing each word of text to those stored in the



database 52, accessing the video movie clip for the word, and causing the media player 51 to play the clip. If a word is not found in the database 52, the text-to-sign language translator 47 breaks the word down into individual letters and plays video clips in the database 52 that show a person finger spelling each letter in the word.

5       The text-to-animation or image translator 48 operates in a similar manner by using each word of text, or groups of words in the text to access an animation/image database 53 that contains animations, cartoons, etc., and/or images that are in some way related to each word or phrase, e.g., pictures of items, etc. For example, the word "cow" can be associated to an image of a cow, while the phrase "the cow jumped over the moon" can be associated with  
10       an animation of a cow jumping over the moon.

      The text-to-phonetic spelling translator 49 accesses a phonetic spelling database 54 that contains text of the phonetic spelling of thousands of different words that can be displayed on the video monitor 44 along with the translated text, animation, still images and signing videos. Finally, the text-to-audio sound translator 50 accesses an audio sound  
15       database 55 that contains brief audio clips of environmental and other sounds that are associated with selected words or phrase in the text stream. For example, the phrase "fire truck" can be associated with the sound of a siren that may be played through the speakers 38, cochlear implant 40 and/or hearing aid 42 simultaneously as the phrase is highlighted on the video display 44. It should be noted in this regard, that the playing of the associated  
20       sound will not occur if the text-to-speech translator 46 is enabled.

      In the preferred embodiment, detected speech is always translated first to text by the speech-to-text translator 45 before it is translated further by any of the remaining translators 46-50. It should be understood, however, that software could also be employed that would receive speech as an input, and generate the signing videos, animation, images, and/or sounds  
25       as output, although some type of word identification procedure would still obviously have to be employed by any such program. Further, as will be discussed in greater detail later, a user of the electronic translator 10 has the option to enable or disable any of the translation functions.

      FIG. 2 shows a sample display screen 56 for the translator 10, and illustrates a number  
30       of windows that are preferably provided. More particularly, the display screen 56 includes a text window 58, an animation/image/phonetic spelling display window 59, a marquee window 60, a signing window 62 and a control window 64. The text window 58 is employed to display text that has been converted by the speech-to-text translator 45. In addition, the

displayed text can be edited and supplemented through use of the keyboard 30. The animation/image/phonetic spelling display window 59 is employed to display images or animations that are associated with highlighted words or phrases in the text window 58, and also to display the phonetic spelling of the words. The marquee window 60 is employed to  
5 display a scrolling version (from right to left) of one line of the text as it is generated by the speech-to-text translator 45, and includes a pair of buttons 65 to increase or decrease the scrolling speed.

The signing window 62 displays the sign language and finger spelling videos, as is shown in greater detail in FIG. 3A. An options menu icon 66 is disposed at the top of the  
10 signing window 62. When selected, as illustrated in FIG. 3B, a drop down menu 68 is generated that allows the signing option to be enabled or disabled. In addition, the play speed of the videos can be adjusted to any one of 8 speeds so that the speed can be synchronized with the text generation speed. It should be noted in this regard that the higher speeds play the signing at speeds that are notably faster than a human can physically sign. As a result,  
15 this feature of the invention is particularly useful in that it enables "speed signing" for individuals that are particularly adept at reading sign language. A pair of buttons 70 is also provided at the top of the signing window 62 that are used to enlarge or reduce the window's size.

The control window 64 facilitates management and organization of the windows, and  
20 the way that they function, and is illustrated in more detail in FIG. 4. A menu bar 72 is provided along the top of the control window 64 for this purpose. As is conventional, the speech-to-text translator 45 can be trained to recognize different voices, and the menu bar 72 can be used to select the identity of the speaker that the speech-to-text translator 45 is to recognize. The identity of the selected speaker is displayed near the bottom of the control  
25 window 64 as illustrated at 74. A text entry box 76 is also disposed in the control window 64 that permits entry, through the keyboard 30, of text or messages for storage or conversion to speech. Through the menu bar 72, other system functions can also be implemented including management of text files, setting of speaking and signing options, and obtaining system help.

Preferably, the windows 58-64 can be selectively docked or unlocked as illustrated in  
30 FIG. 5, and can be selectively enabled or disabled (the animation/image/phonetic spelling display window 59 is disabled in FIG. 5). Docking allows a user to arrange the four windows in any way that the user prefers. To facilitate this function, a plurality of latch icons 78 are employed to indicate whether the windows are locked to one another, and can be "clicked on"

with the mouse 31 to be latched or unlatched. When the windows are unlatched, they can be freely moved relative to one another.

FIGs. 6A-6H are flowcharts that illustrate the preferred method of operation of the electronic translator 10. With reference first to FIG. 6A, there are two distinct flow streams to the method, one for input of text from the keyboard 30, and a second for input from other of the peripheral devices, such as from the microphone 32. First, if input is determined at step 100 to be received from the keyboard 30, the entered text is displayed on the video display 44 at step 102, and is recorded or stored at step 104. Next, referencing FIG. 6F, if the telephone 36 is attached and activated at step 106, then the text is converted to speech at step 108, and sent to the cell phone 36 at step 110 via the PCMCIA card 23 or the sound card 22. If the cell phone 36 is not attached or not activated and the keyboard input is to be spoken at step 112, then the right channel of the sound output is turned on at step 114. Next, the text is spoken at step 116 and played through the speakers 38, then the right channel of the sound output is turned off at step 118. It should be noted in this regard that the right channel of the sound card 22 is activated only when keyboard entered text is to be translated and spoken to avoid having the speakers 38 play speech that has just been spoken into the microphone 32. If the keyboard input is not to be spoken at step 112, then the software goes into a wait state at step 120 to wait for the next data input.

Returning to FIG. 6A, if the input from the data input means was determined at step 100 not to be a keyboard input, then the input is assumed to be speech from one of the audio input devices. The speech input is converted to text at step 122, put into a buffer at step 124, and applied to multiple branches of the method flow at this point. If it is determined at step 126 that the text display box is enabled, then the text is displayed via the video display 44 at step 128. Also, if the marquee option is enabled at step 130, then the text is displayed marquee style in the marquee window 60 on the video display 44 at step 132.

The third branch of the method flow is employed for the conversion of the text into sign language, animation or images and environmental sounds. To perform these translations, the phrases of text in the text buffer must first be parsed into separate words at step 134, the punctuation is removed at step 136, and the words are put into a queue at step 137. With reference to FIG. 6B, multiple queries are then performed to determine which translation options are enabled, and then carry out the appropriate steps for each. If the "speak with signing" option is determined to be on at step 138, then the process continues as illustrated in FIG. 6C. At step 140, it is determined whether there is a word to speak. If so,

the word is spoken at step 142 through use of the text-to-speech translator 46 and one or more of the audio devices. In addition, the word is highlighted on the video display 44 at step 144 as it is spoken. Once it is determined that there are no more words left to be spoken, the process goes into a wait state at step 146.

5           Referencing again FIG. 6B and also referencing FIG. 6D, if the signing option is determined to be on at step 148, the first word in the queue is retrieved at step 150, and the word is displayed and highlighted at step 152. A search is performed against the signing database 52 for that word at step 154. If it is determined at step 156 that the word is in the signing database 52, then the directory location and filename for the video for that particular  
10       word is retrieved at step 158. If it is determined at step 160 that "speak with signing" is turned on, then the word is spoken at step 162 via the sound card 22 to the appropriate ones of the audio devices. If a video was found at step 164, then the retrieved directory location and filename path is assigned to the media-player 51 at step 166. The playback speed of the video is adjusted according to the user's selection at step 168, and the media-player 51 is  
15       given the command to play the movie of the human performance of the sign language of the word at step 170.

          If the word is not found in the signing database 52 at step 156, and it is determined at step 172 that the word has an ending that can be removed, then the ending is stripped off at step 174 and the database is searched again. If the word does not have an ending that can be  
20       removed at step 172 and the finger spelling option is turned on at step 176, the process continues as illustrated in FIG. 6E. The number of letters in the word is counted at step 178 and the next letter is retrieved at step 180. The directory location and filename for the video for that particular letter is retrieved at step 182. The current letter and word is then displayed at step 184. Next, the retrieved directory location and filename path is assigned to the media  
25       player 51 at step 186, and the playback speed is optionally adjusted at step 188. The media player 51 is given the command to play the movie of the human performance of the finger spelled letter at step 190. If it is determined at step 192 that there is another letter in the word, then the process returns to step 180, and the next letter is retrieved. Otherwise, the "set video location found" switch is set to "No" at step 194 so that the "was video found"  
30       indicator returns "No" and the process returns to the "was video location found" step 164 in FIG. 6D.

          Once the video movies of the word signing or finger spelling of the letters have been played, the process continues at step 196 where it checks to see if the signing option is on. If

the signing option is off, then all the words are removed from the queue at step 198, the current word display is cleared at step 200, and the software goes into a wait state at step 202. If the signing option is on, then the first word is removed from the queue at step 204 and the queue is checked for another word at step 206. If there is another word in the queue, then the  
5 program returns to step 150 and the first word in the queue is retrieved as discussed previously. If there is not another word in the queue, then the current word is cleared from the display at step 200 and the software goes into the wait state at step 202.

Returning once again to FIG. 6B, and also referencing FIG. 6G, if the "show images" option is determined to be on at step 236, then the first word in the queue is retrieved at step  
10 238, and the phonetic spelling of the word is displayed at step 240. A search is performed against the animation/image database 53 for that word (or phrase) at step 242. If it is determined at step 244 that the word or phrase is in the database 53, then the directory location and filename for the image or animation for that particular word or phrase is retrieved at step 246. If it is determined at step 248 that "speak with signing" is turned on,  
15 then the word is spoken at step 250 via the sound card 22 to the appropriate ones of the audio devices. If it is determined at step 252 that an image location was found, then the image is displayed at step 254.

If the word is not found in the image database 53 at step 244, and it is determined at step 256 that the word has an ending that can be removed, then the ending is stripped off at  
20 step 258 and the database is searched again. Once the image display process is finished, the first word in the queue is removed at step 260, and a query is made at step 262 whether there is another word in the queue. If so, the process returns to step 238 for processing of the next word. If not, the current word is cleared from the display at step 264, and the process goes into the wait state at step 266.

25 The last translation function that can be optionally employed with the electronic translator 10 is the conversion of words or phrases to related environmental sounds by the text-to-audio sound translator 50. If this option is determined to be enabled at step 268 as illustrated in FIG. 6B, the process continues as illustrated in FIG. 6H. At step 270, the first word in the queue is retrieved, and a search is performed against the audio sounds database  
30 55 for that word at step 272. If it is determined at step 274 that the word is in the database 55, then the directory location and filename for the sound file for that particular word is retrieved at step 276. If it is determined at step 278 that the sound file location was found,

then the directory location and name of the sound file are assigned to the media player 51 at step 280, and the sound associated with the word is played at step 282.

5 If the word is not found in the audio sounds database 55 at step 274, and it is determined at step 284 that the word has an ending that can be removed, then the ending is stripped off at step 286 and the database 55 is searched again. Once the search and sound play process is finished, the first word in the queue is removed at step 288, and a query is made at step 290 whether there is another word in the queue. If so, the process returns to step 268 for processing of the next word. If not, the current word is cleared from the display at step 292, and the process goes into the wait state at step 294.

10 In conclusion, the present invention represents a tremendous step forward in the field of communication and learning assistance for hearing impaired individuals. The generation of simultaneous multiple data streams in different formats and in real time provides a powerful learning and communications tool that heretofore has not been available. It should be noted that while the invention is particularly suited for use with hearing impaired  
15 individuals, the invention is also applicable to the blind deaf, the hearing mute and those individuals defined as less than average learners suffering from a variety of physical, mental or emotional disabilities. The invention is also applicable to individuals with no related disabilities who nevertheless have an interest in learning sign language or improving reading skills.

20 Although the invention has been disclosed in terms of a preferred embodiment and variations thereon, it will be understood that numerous additional modifications and variations could be made thereto without departing from the scope of the invention as defined in the following claims.

CLAIMS

1. An electronic translator for translating speech into multiple forms comprising:

a) a receiver for generating a speech signal in response to a person speaking;

b) a processor, said processor including:

5                   1) a speech-to-text translator for converting said input speech signal into a text data stream; and

                  2) a text-to-speech translator for converting said text data stream into an audio signal;

                  c) a video display interfaced to said processor for displaying text that corresponds to  
10 said text data stream; and

                  d) an audio output device interfaced to said processor for receiving said audio signal and generating a speech sound simultaneously with the display of said text.

2. The electronic translator of claim 1, further including a text-to-sign language  
15 translator for converting said text data stream into a video image of a person using sign language to sign words in said data stream, and displaying said video image simultaneously with said text on said display.

3. The electronic translator of claim 2, wherein said text-to-sign language translator  
20 includes a database containing a plurality of video clips of a person signing words, each of said video clips showing the signing for a particular word; and means responsive to said text data stream for detecting words therein, sequentially retrieving video clips in said database showing signing of said words, and playing said video clips sequentially on said video display.

25  
4. The electronic translator of claim 3, wherein said text-to-sign language translator database further includes a plurality of finger spelling video movies for signing individual letters of the alphabet, and said text-to-sign language translator further includes means for determining that a word in said text data stream does not have a corresponding signing video  
30 in said database, and accessing said finger spelling videos to display sign language spelling of said word.

5. The electronic translator of claim 3, wherein said text-to-sign language translator includes speed selection means for adjusting the play speed of said video clips to enable synchronization of said video clips with said text as they are simultaneously displayed on said video display.

5

6. The electronic translator of claim 3, further including highlighting means for highlighting words in text being displayed on said video display by said speech-to-text translator as said highlighted words are being signed by said text-to-sign language translator, and generated by said audio output device.

10

7. The electronic translator of claim 6, wherein said audio output device comprises a cochlear implant, whereby, a hearing impaired person can identify a sound generated by said cochlear implant as corresponding to said highlighted words.

15

8. The electronic translator of claim 1, further comprising a text-to-image translator for translating words in said text into one or more graphical images that are associated with said words, and displaying said images simultaneously with said text on said video display.

20

9. The electronic translator of claim 1, wherein said processor further includes means for simultaneously displaying text on said display in a first scrolling marquee format and in a second static format.

25

10. The electronic translator of claim 1, wherein said processor further includes means for displaying a plurality of information windows on said video display, including at least a first window for displaying translated text and a second window for displaying control functions.

30

11. The electronic translator of claim 10, wherein said processor further includes means for selectively latching and unlatching said windows to and from one another to facilitate arrangement of said windows on said display either as a group, or individually.

12. The electronic translator of claim 1, wherein said processor further includes a text-to-phonetic spelling translator, including a phonetic spelling database, for translating



words in said text data stream into text of phonetic spelling of said words, and displaying said phonetic spelling text on said video display.

13. The electronic translator of claim 1, wherein said processor further includes a  
5 text-to-audio sound translator, including an audio sound database, for translating words in said text data stream into environmental sounds that are related to said words, and playing said audio sounds on said audio output device.

14. An electronic translator for translating text into multiple forms comprising:  
10 a) means for receiving a text data stream to be translated;  
b) a processor, said processor including a text-to-sign language translator for  
converting said text data stream into a video image of a person using sign language to sign words in said data stream; and  
c) a video display interfaced to said processor for displaying said text stream  
15 simultaneously with said video image.

15. The electronic translator of claim 14, wherein said processor further includes a text-to-speech translator for converting said text stream into an audio signal, and said translator further comprises an audio output device for receiving said audio signal, and  
20 generating audio sounds in response thereto.

16. The electronic translator of claim 14, wherein said text-to-sign language translator includes a database containing a plurality of video clips of a person signing words, each of said video clips showing the signing for a particular word; and means responsive to  
25 said text data stream for detecting words therein, sequentially retrieving video clips in said database showing signing of said words, and playing said video clips sequentially on said video display.

17. The electronic translator of claim 16, wherein said text-to-sign language  
30 translator database further includes a plurality of finger spelling video movies for signing individual letters of the alphabet, and said text-to-sign language translator further includes means for determining that a word in said text data stream does not have a corresponding

signing video in said database, and accessing said finger spelling videos to display sign language spelling of said word.

18. The electronic translator of claim 16, wherein said text-to-sign language  
5 translator includes speed selection means for adjusting the play speed of said video clips to enable synchronization of said video clips with said text as they are simultaneously displayed on said video display.

19. The electronic translator of claim 14, further comprising a text-to-image  
10 translator for translating said text into one or more images that are associated with said text, and displaying said images simultaneously with said text on said video display.

20. The electronic translator of claim 14, wherein said processor further includes  
15 means for simultaneously displaying text on said display in a first scrolling marquee format and in a second static format.

21. The electronic translator of claim 14, wherein said processor further includes  
means responsive to inputs from said input device for establishing a telephone connection  
with a telephone, transmitting said audio signal over a transmitter in said telephone, and  
20 receiving a speech input signal from a receiver in said telephone for translation by said speech-to-text translator and display on said video display monitor.

22. The electronic translator of claim 14, wherein said processor further includes  
means for displaying a plurality of information windows on said video display, including at  
25 least a first window for displaying translated text and a second window for displaying control functions.

23. The electronic translator of claim 22, wherein said processor further includes  
means for selectively latching and unlatching said windows to and from one another to  
30 facilitate arrangement of said windows on said display either as a group, or individually.

24. The electronic translator of claim 14, wherein said processor further includes a  
text-to-phonetic spelling translator, including a phonetic spelling database, for translating

words in said text data stream into text of phonetic spelling of said words, and displaying said phonetic spelling text on said video display.

25. The electronic translator of claim 14, wherein said processor further includes a  
5 text-to-audio sound translator, including an audio sound database, for translating words in said text data stream into environmental sounds that are related to said words, and playing said audio sounds on an audio output device.

26. A method for translating speech to multiple formats comprising the steps of:  
10 a) generating a speech signal in response to a person speaking;  
b) translating said speech signal into a text data stream;  
c) translating said text data stream into an audio signal;  
d) driving an audio output device with said audio signal; and  
e) displaying text that corresponds to said text data stream on a video display.

15 27. The method of claim 26, wherein the step of driving an audio output device with said audio signal further comprises driving a cochlear implant with said audio signal.

28. The method of claim 26, further including the steps of translating said text data  
20 stream into a video image of a person using sign language to sign words in said data stream, and displaying said video image simultaneously with said text on said display.

29. The method of claim 28, wherein said step of translating said text data stream  
25 into a video image of a person using sign language further comprises the steps of accessing a database containing a plurality of video clips of a person signing words, each of said video clips showing the signing for a particular word; sequentially retrieving video clips in said database showing signing of said words, and playing said video clips sequentially on said video display.

30 30. The method of claim 29, wherein said database further includes a plurality of finger spelling video movies for signing individual letters of the alphabet, and said method further comprise the steps of determining that a word in said text data stream does not have a

corresponding signing video in said database, sequentially retrieving finger spelling videos corresponding to the spelling of said word, and displaying said finger spelling videos.

31. The method of claim 29, further comprising the step of adjusting a playback  
5 speed of said video clips to synchronize of said video clips with said text as they are simultaneously displayed on said video display.

32. The method of claim 26, further including the step of highlighting words in text  
being displayed on said video display as said highlighted words are being signed, and  
10 generated by said audio output device;

whereby, a hearing impaired person with a cochlear implant can identify a sound generated by a cochlear implant as corresponding to said highlighted words.

33. The method of claim 32, further comprising the step of displaying a phonetic  
15 spelling representation of a word as it is highlighted.

34. The method of claim 32, further comprising the step of playing an audio sound that is related to a highlighted word.

20 35. The method of claim 26, further comprising the step of translating said text into one or more images that are associated with said text, and displaying said images simultaneously with said text on said video display.

25 36. A method for translating a text data stream into multiple formats comprising the steps of:

- a) generating a text data stream;
- c) translating said text data stream into an audio signal;
- 30 d) driving an audio output device with said audio signal; and
- e) displaying text that corresponds to said text data stream on a video display.

37. The method of claim 36, wherein the step of driving an audio output device with said audio signal further comprises driving a cochlear implant with said audio signal.

38. The method of claim 36, further including the steps of translating said text data stream into a video image of a person using sign language to sign words in said data stream, and displaying said video image simultaneously with said text on said display.

39. The method of claim 38, wherein said step of translating said text data stream into a video image of a person using sign language further comprises the steps of accessing a database containing a plurality of video clips of a person signing words, each of said video clips showing the signing for a particular word; sequentially retrieving video clips in said database showing signing of said words, and playing said video clips sequentially on said video display.

40. The method of claim 39, wherein said database further includes a plurality of finger spelling video movies for signing individual letters of the alphabet, and said method further comprise the steps of determining that a word in said text data stream does not have a corresponding signing video in said database, sequentially retrieving finger spelling videos corresponding to the spelling of said word, and displaying said finger spelling videos.

41. The method of claim 39, further comprising the step of adjusting a playback speed of said video clips to synchronize of said video clips with said text as they are simultaneously displayed on said video display.

42. The method of claim 36, further including the step of highlighting words in text being displayed on said video display as said highlighted words are being generated by said audio output device;

whereby, a hearing impaired person with a cochlear implant can identify a sound generated by a cochlear implant as corresponding to said highlighted words.

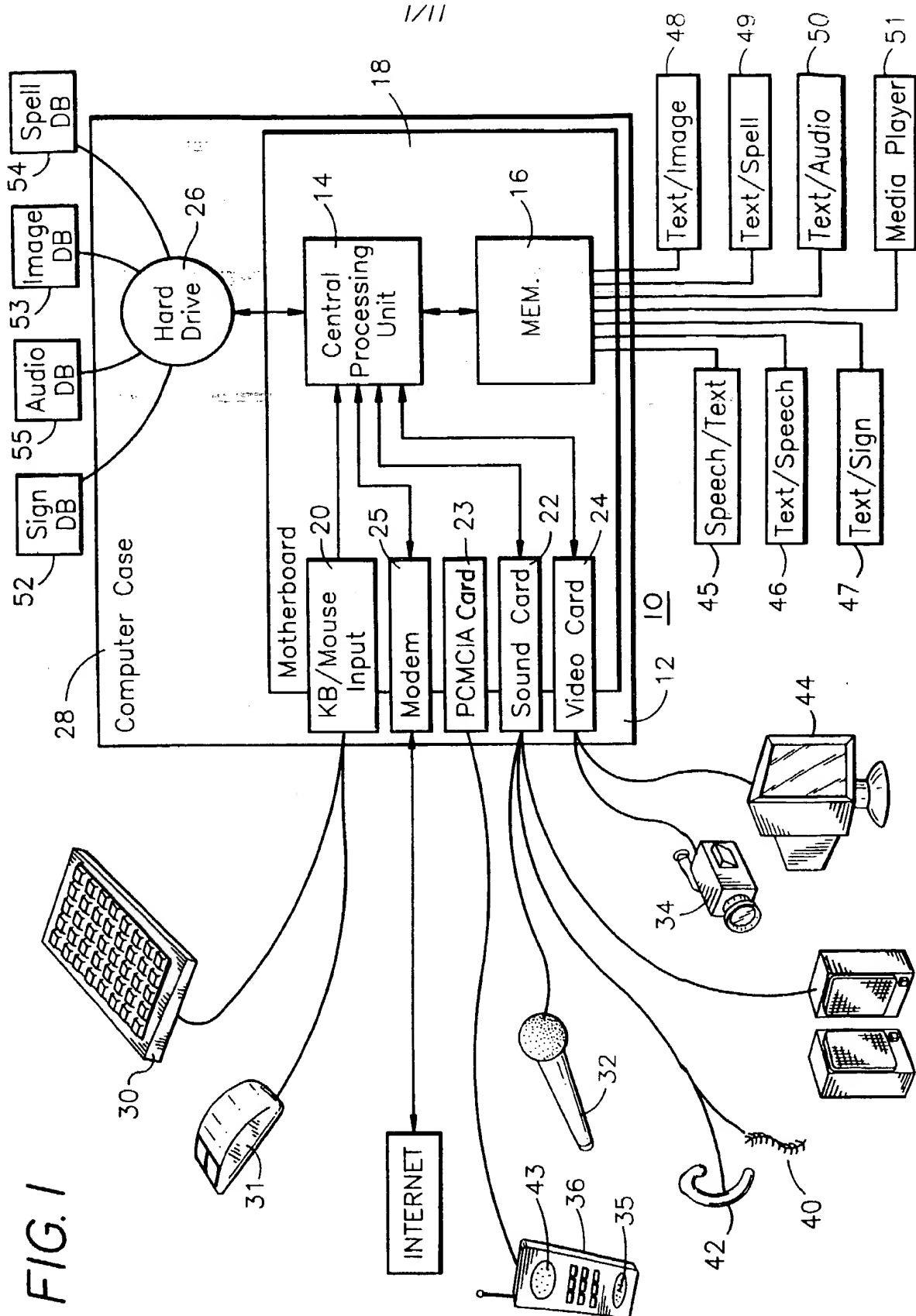
43. The method of claim 42, further comprising the step of displaying a phonetic spelling representation of a word as it is highlighted.

44. The method of claim 42, further comprising the step of playing an audio sound that is related to a highlighted word.

45. The method of claim 36, further comprising the step of translating said text into one or more images that are associated with said text, and displaying said images simultaneously with said text on said video display.

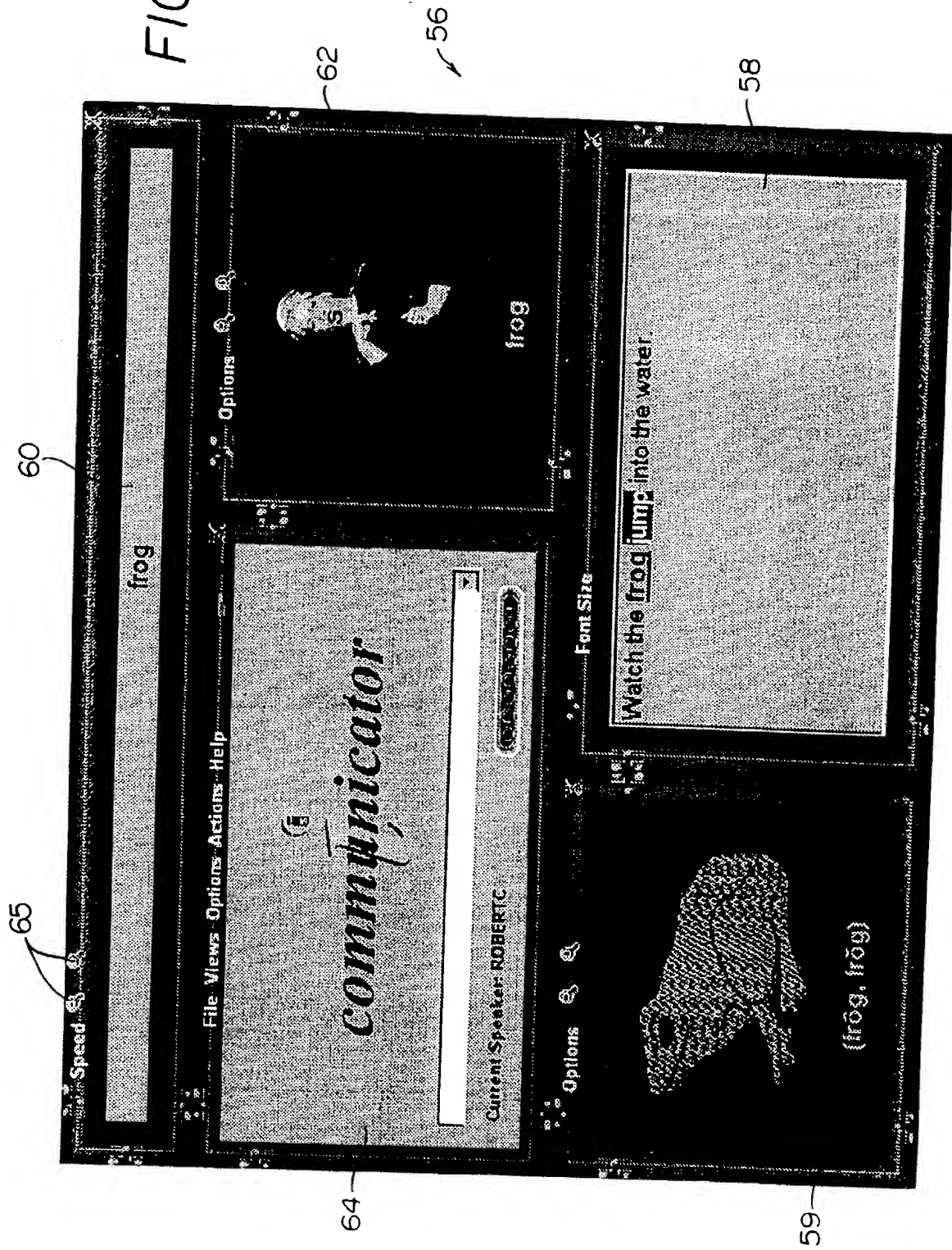
46. The method of claim 36, further comprising the steps of establishing a telephone connection with a telephone, transmitting said audio signal over a transmitter in said telephone, receiving a second speech input signal from a receiver in said telephone, and translating said second speech signal into a second text data stream, and displaying said second text data stream on said video display.

15

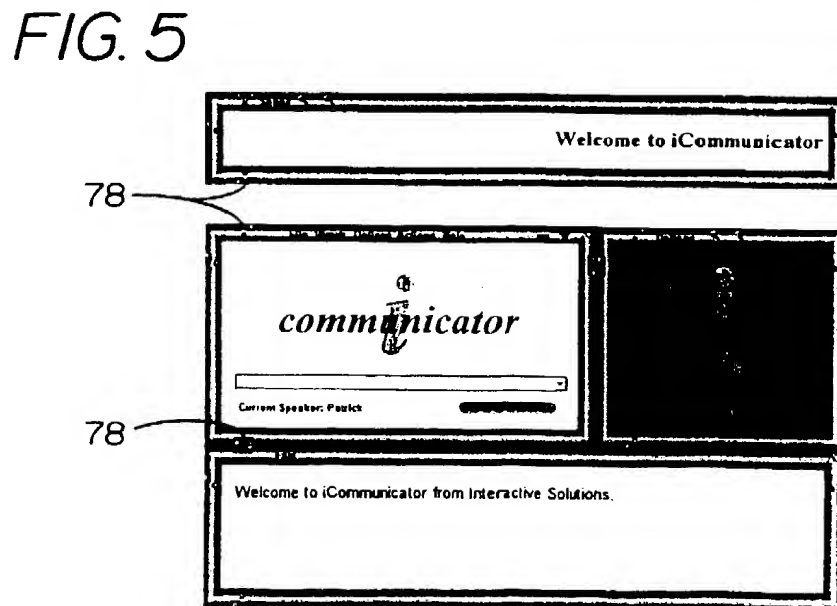
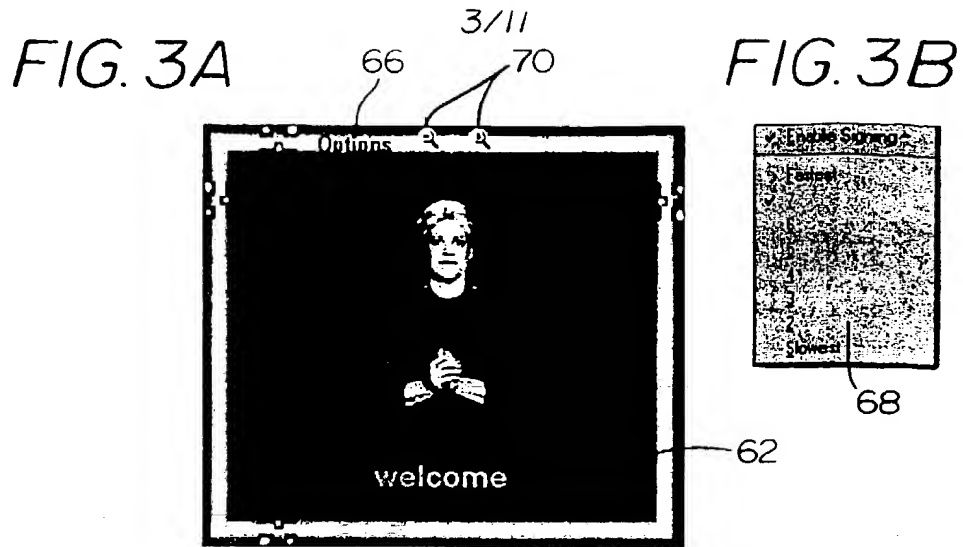


2/11

FIG. 2

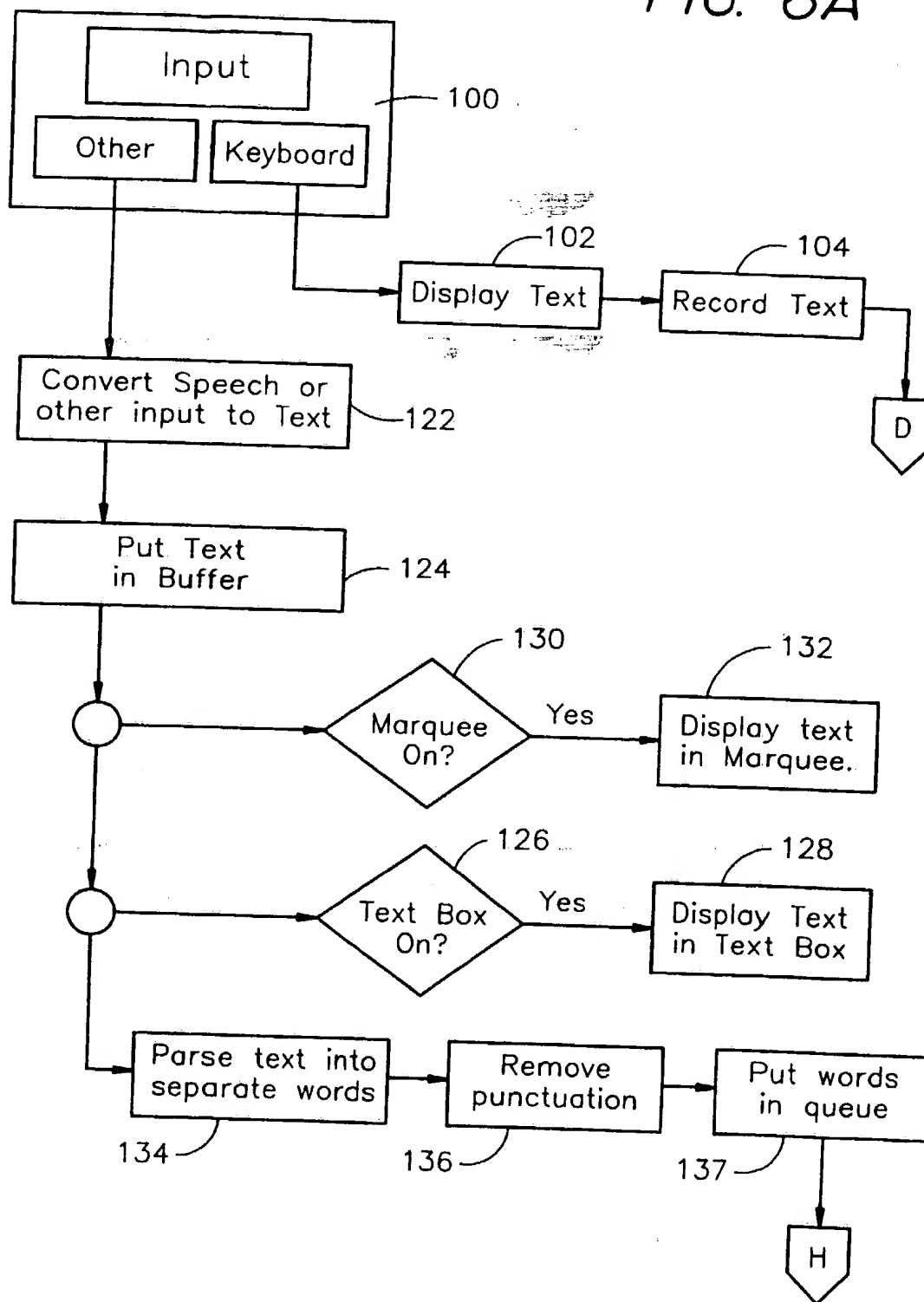






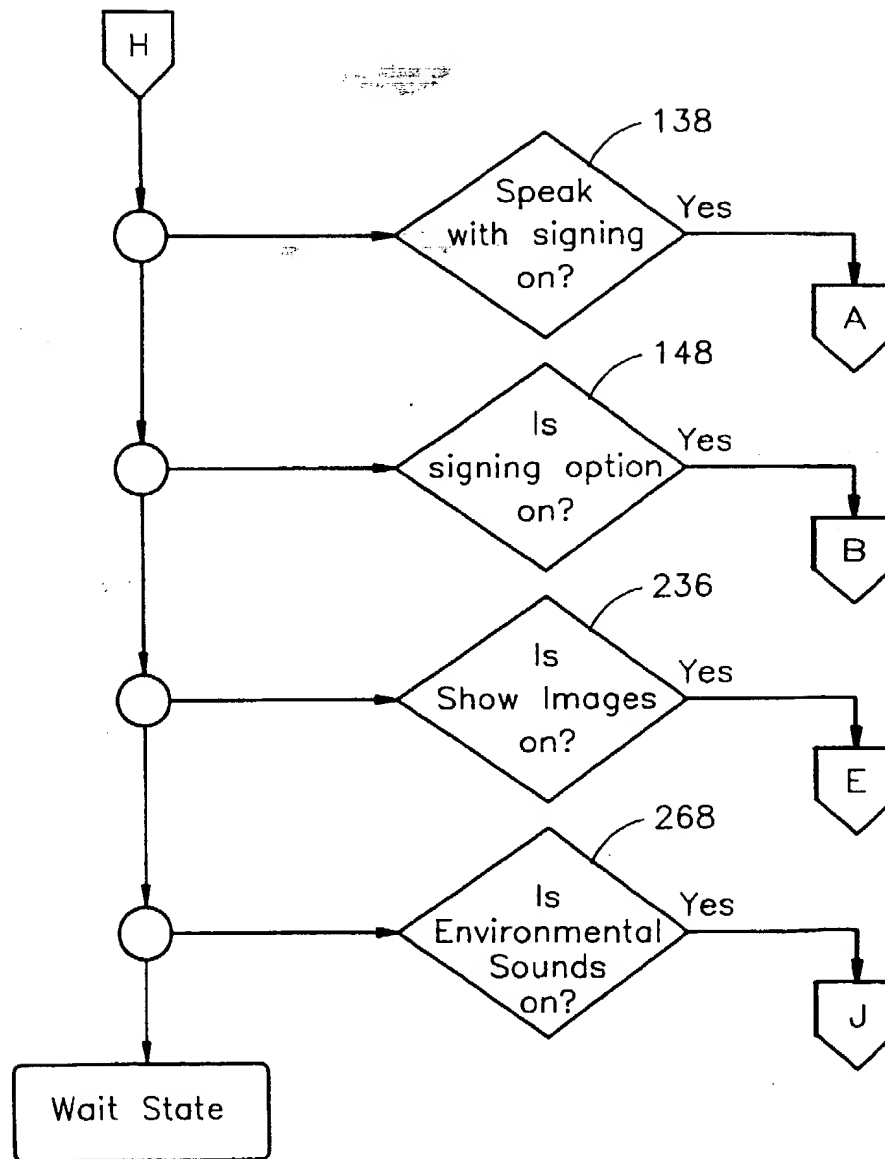
4/11

FIG. 6A



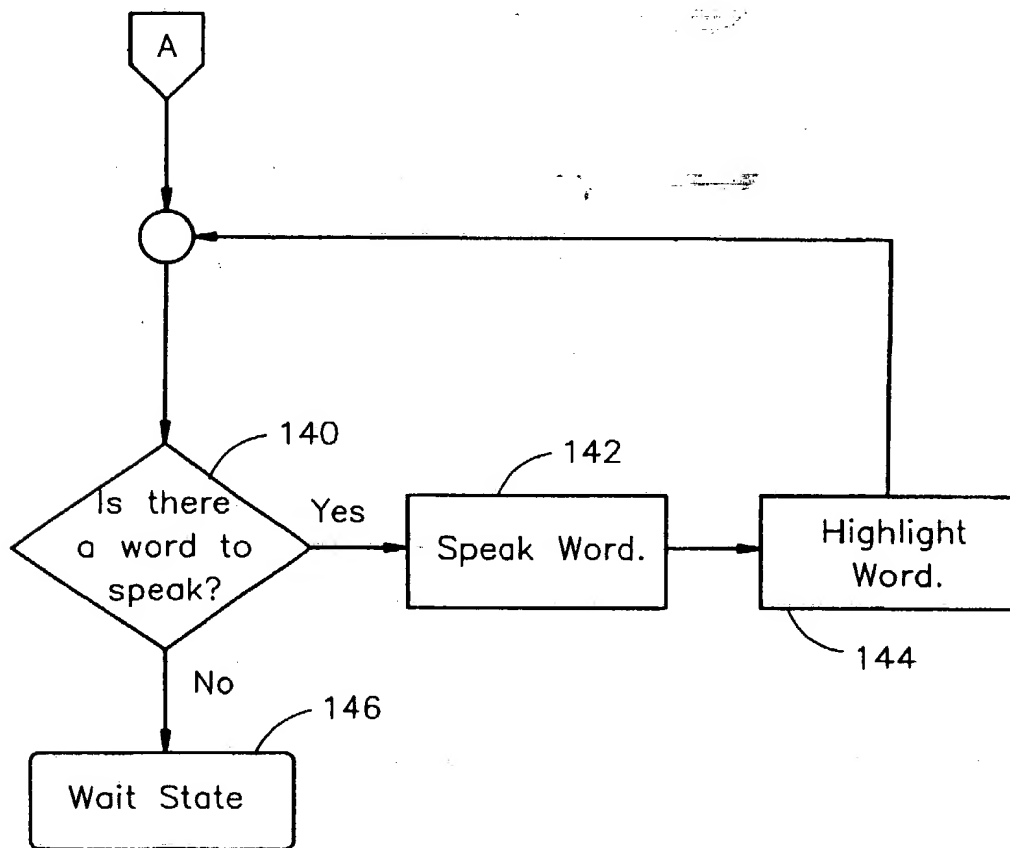
5/11

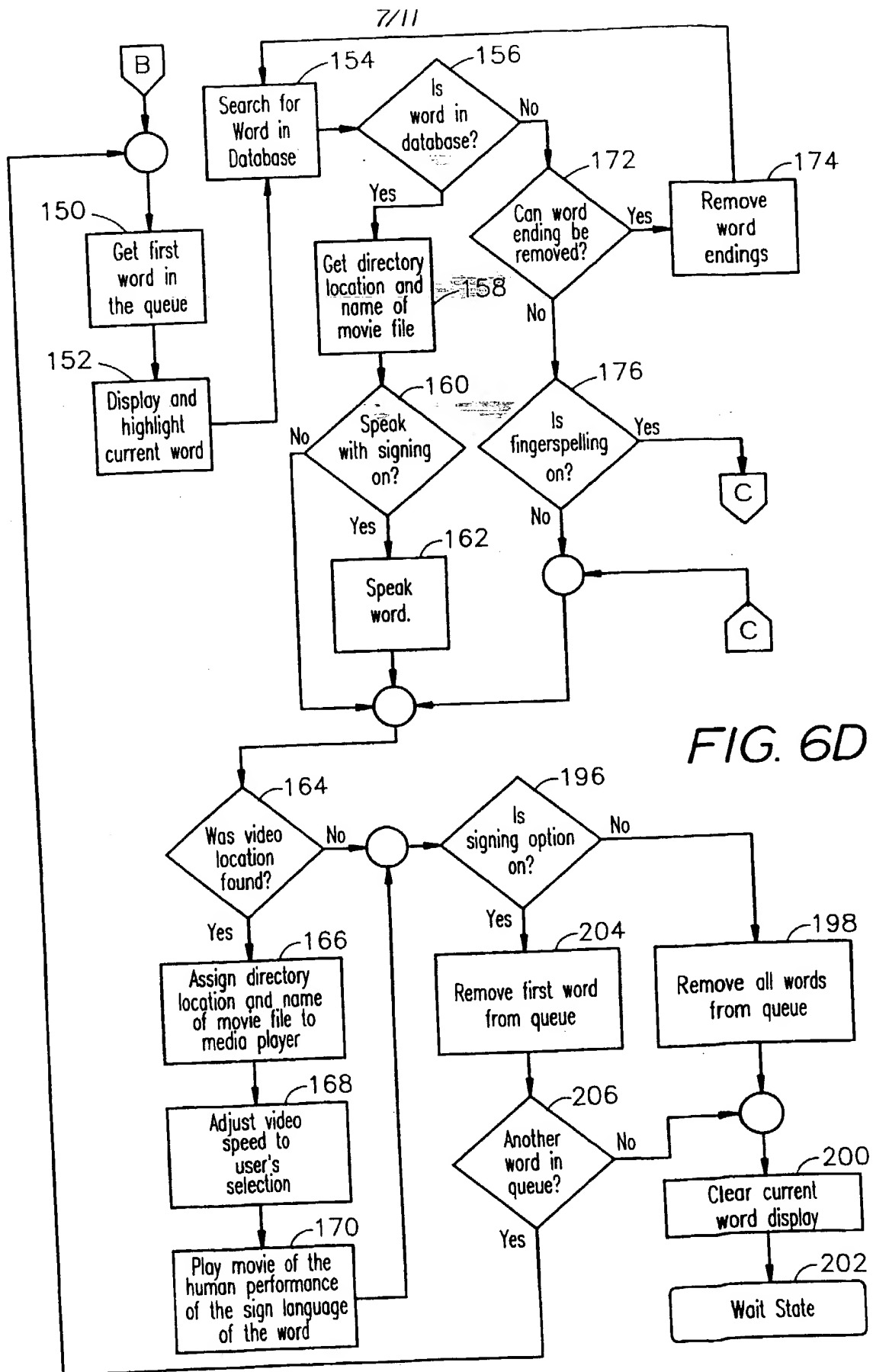
FIG. 6B



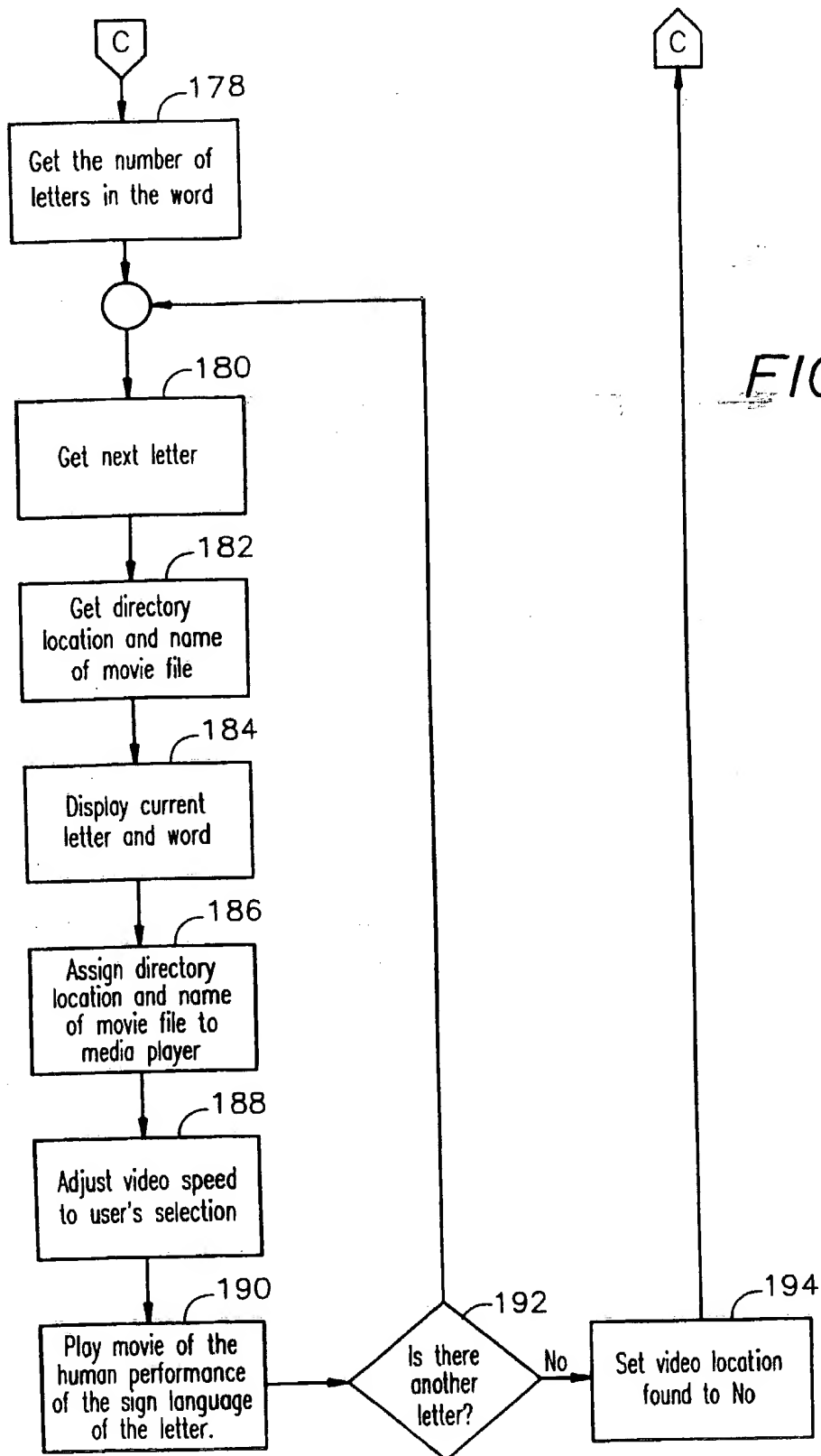
6/11

FIG. 6C



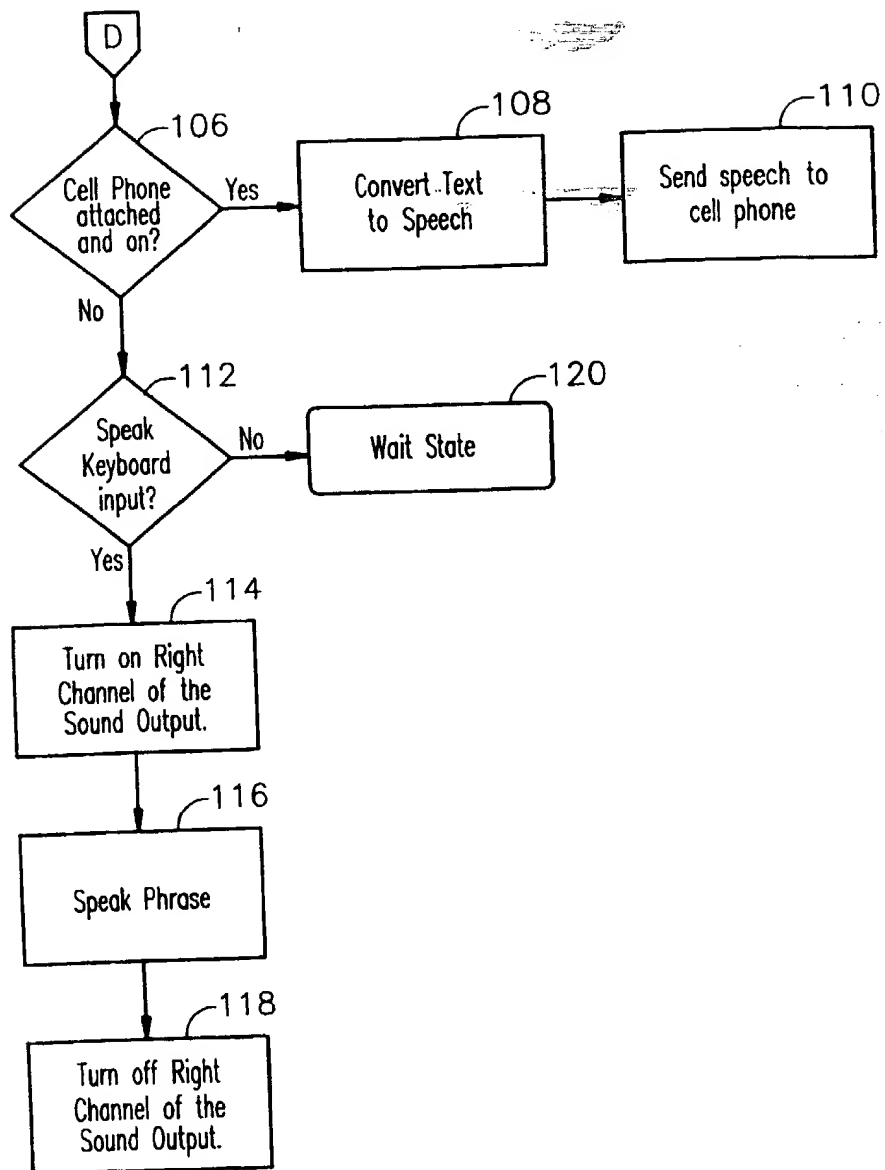


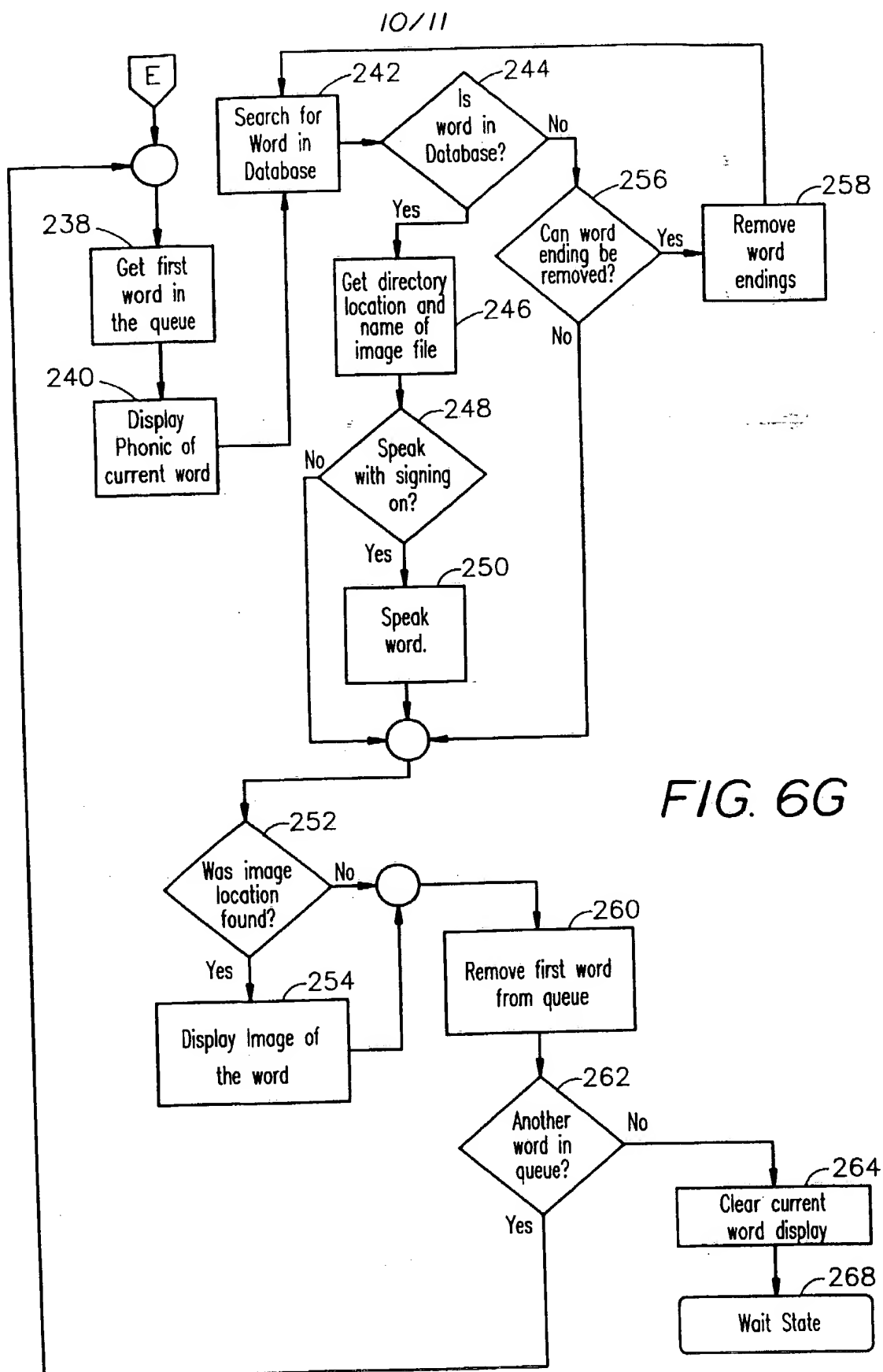
8/11



9/11

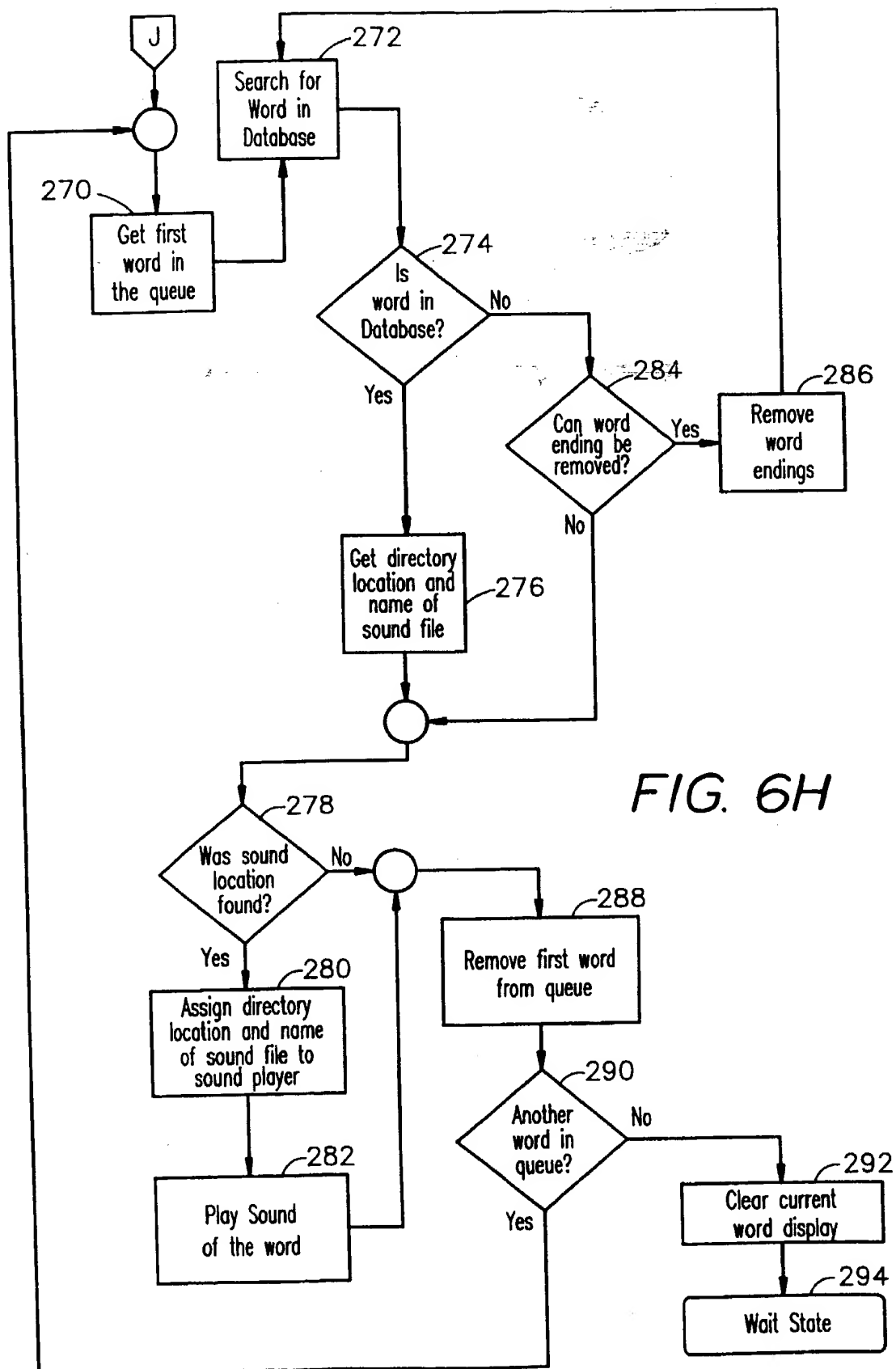
FIG. 6F







11/11



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/32452

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(7) : G10L 13/08, 15/26, 21/06 US CL : 704/235, 260, 271 According to International Patent Classification (IPC) or to both national classification and IPC														
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. : 704/235, 260, 271  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) USPAT, EPO, JPO, DWPI, TDBD, IEEE Xplore														
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>														
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.												
Y	US 5,659,764 A (SAKIYAMA et al) 19 August 1997, Fig. 22 and 23 and accompanying discussion.	1-46												
Y	US 5,890,120 A (HASKELL et al) 30 March 1999, col. 2, lines 19-65; and col. 3, lines 1-2.	1-46												
Y,E	US 6,181,351 B1 (MERRILL et al) 30 January 2001, Abstract, Figure 4 and accompanying discussion.	1-46												
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.														
<table border="0"><tr><td>* Special categories of cited documents:</td><td>*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</td></tr><tr><td>*A* document defining the general state of the art which is not considered to be of particular relevance</td><td>*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</td></tr><tr><td>*B* earlier document published on or after the international filing date</td><td>*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</td></tr><tr><td>*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</td><td>*G* document member of the same patent family</td></tr><tr><td>*O* document referring to an oral disclosure, use, exhibition or other means</td><td></td></tr><tr><td>*P* document published prior to the international filing date but later than the priority date claimed</td><td></td></tr></table>			* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family	*O* document referring to an oral disclosure, use, exhibition or other means		*P* document published prior to the international filing date but later than the priority date claimed	
* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention													
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone													
*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art													
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family													
*O* document referring to an oral disclosure, use, exhibition or other means														
*P* document published prior to the international filing date but later than the priority date claimed														
Date of the actual completion of the international search 13 FEBRUARY 2001		Date of mailing of the international search report 04 APR 2001												
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer TALIVALDIS IVARS SMITS Telephone No. (703) 306-3011												